

# Application of reinforcement learning in mechatronic systems

First-Name Last-Name1, First-Name Last-Name2, First-Name Last-Name3

1 Affiliation, E-mail:

2 Affiliation, E-mail:

3 Affiliation, E-mail:

\*Corresponding E-mail:

**Abstract:** 近年來硬體技術、軟體、自動導等技術快速發展起來，再次帶起機器學習的發展，促使機器學習與各領域結合的應用越來越廣泛，在機電系統採用強化學習是為了讓機電系統的控制達到最佳化。本研究利用強化學習優化冰球機的對打系統，並測試將相同的訓練算法能在 2D 環境上進行訓練和使用，將相同演算法套用到 3D 模擬環境進行訓練與使用，以驗證算法若將既有機電系統簡化成 2D 所測試的算法套用到 3D 模擬環境的可行性。將實體冰球機的機電系統簡化後導入 CoppeliaSim 模擬環境透過 Remote API 控制環境中的冰球機移動，OpenCV 來處理影像提供強化學習訓練和訓練後實際控制的輸入，強化學習訓練利用 OpenAI Gym 的 Pong Game 測試適合的訓練參數，再將算法套用到 CoppeliaSim 的場景中進行訓練。

**Keywords:** 類神經網路、強化學習、CoppeliaSim、OpenAI Gym

## Introduction

由於機器學習的訓練需要透過長時間的訓練，測試合適的機器學習算法和參數需要長時間的測試與訓練才能得到適用的演算法，若以實體的機電系統進行訓練需要投入大量的金錢與時間成本，採取虛擬訓練的方式：在相同時間下可以訓練的次數較多實體環境中多，改善訓練時間冗長和組裝實體系統的成本。測試在 2D 環境中的演算法在 3D 環境下的可行性並評估演算法若套用到實體系統可行性。

### 相關文獻

透過 OpenAI Gym 內建編譯的 Pong 遊戲環境當作 2D 的訓練環境，3D 則是透過 CoppeliaSim 來模擬訓練機電系統的運作狀況，訓練演算法選用強化學習結合類神經網路的學習方式，其中使用 score function gradient 當作主要算法。

### 概述工作的主要結果

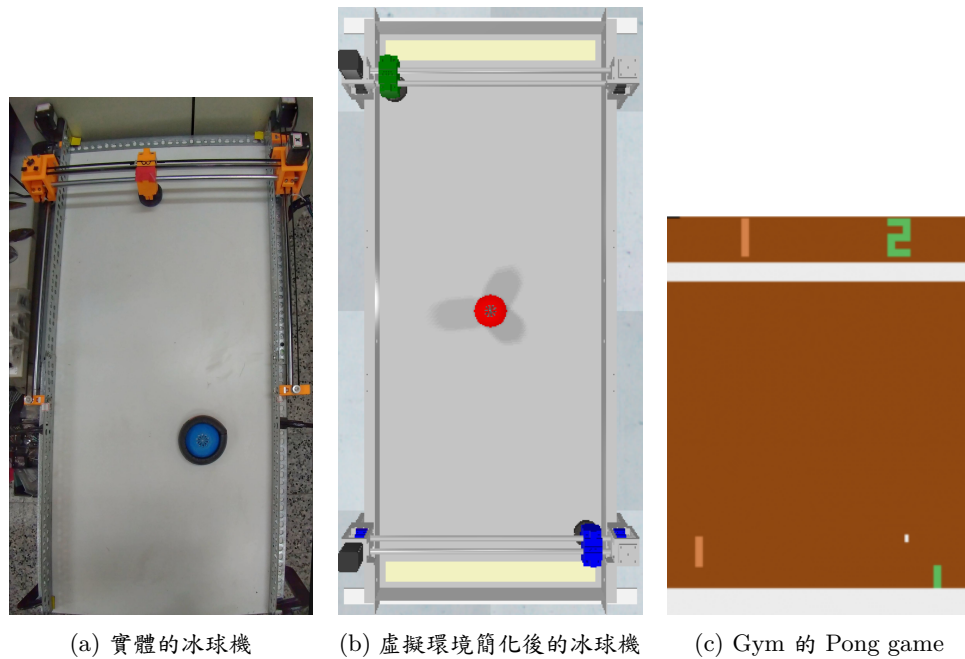


Figure 1: 冰球機與訓練環境

## Methodology

強化學習是環境和 agent 互動並互相影響著，以獎勵的方式鼓勵學習，並且不需要特別教導。類神經網路則是透過神經元之間的非線性與權重產生記憶性並學習，以 back-propagation 的方式修正權重與偏差達到進步的效果。結合類神經網路和強化學習這兩種算法，使訓練不需特別教導並以獎懲的方式鼓勵最大化得分，並在固定次數訓練後 back-propagation 修正權重偏差。

參考資料

解釋研究問題

描述研究框架

將對打系統的基本移動控制藉由強化學習來達到最佳化，透過 Gym 測試演算法可行性再進一步運用 CoppeilaSim 進行 3D 模擬環境訓練對打系統，後續可藉由冰球機結合網路攝影機將整個對打系統實際運用。

應用的方法

研究問題與理論和實踐相關

為什麼選擇的方法解決該問題

## Findings

論點和發現

## Conclusions

推斷結果的原則和概括

工作的任何例外、問題或限制

工作的理論和或實踐意義

得出的結論和建議

## Acknowledgements

致謝

## References

- [1] <http://blog.shakirm.com/2015/11/machine-learning-trick-of-the-day-5-log-derivative-trick/>